DOCUMENT RESUME

ED 478 979 TM 035 124

AUTHOR	Kim, Jimmy
TITLE	The Initial Response to the Accountability Requirements in the No Child Left Behind Act: A Case Study of Virginia and Georgia.
PUB DATE	2003-04-21
NOTE	44p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25, 2003).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	*Accountability; Case Studies; *Civil Rights; Educational Policy; Elementary Secondary Education; Federal Legislation; *Political Influences; Politics of Education; State Programs; *State Standards
IDENTIFIERS	*Georgia; No Child Left Behind Act 2001; *Virginia

ABSTRACT

This study had two broad goals: to discuss state politics and accountability policies in Virginia and Georgia before the No Child Left Behind (NCLB) Act and to analyze the civil rights implications of NCLB's accountability requirements for future policy decisions. Part 1 describes previous policies and governance arrangements and how these factors affect state capacity to meet NCLB's accountability requirements. This description is based on state policy documents, newspaper articles, and interviews with state officials to provide a short narrative history of Virginia and Georgia efforts to construct an accountability system in the 1990s. Part 2 analyzes the implications of NCLB's accountability requirements for state policy decisions, especially with respect to establishing adequate yearly progress (AYP) starting points and improvement goals. In response to NCLB. Both states adopted the 12-year timeline for having all students become proficient on the state assessment. As a result, both Virginia and Georgia set initial pass rates and AYP targets for bringing all students to the proficient level. The study used achievement results from state tests and the National Assessment of Educational Progress (NAEP) to compare proficient levels on state assessments in Virginia and Georgia with the NAEP performance standards in grade 8 mathematics. Findings show that proficient performance on both state tests corresponds to the "basic" level of performance on the NAEP for all major subgroups. Although it may be reasonable for states to improve the percentage of students meeting the NAEP basic level on an annual basis, the report highlights several assumptions that would have to be met in order to reach these goals. The study also discusses the civil rights implications of the racial subgroup rules, which may disadvantage racially diverse schools unfairly. The initial response of these two states suggests that the implementation of NCLB prescriptive accountability requirements will depend on the cohesiveness of state accountability policies before the enactment of NCLB and coordinated policy making between the State Board of Education, the Sate Education Agency, and state political leaders. (Contains 4 tables, 5 figures, and 32 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

TM TMTR 0161

ED 478 979

THE INITIAL RESPONSE TO THE ACCOUNTABILITY REQUIREMENTS IN THE NO CHILD LEFT BEHIND ACT: A CASE STUDY OF VIRGINIA AND GEORGIA

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
J. Kim
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

Jimmy Kim
The Civil Rights Project
Harvard University

Working Draft

TM035124

Prepared for the annual meeting of the American Educational Research Association, Chicago, IL, April 21-25, 2003.

BEST COPY AVAILABLE

TABLE OF CONTENTS

List of Tables 2

List of Figures 3

Introduction: NCLB and State Responsibilities 4
State Implementation of Standards-Based Accountability 5
Lessons Learned from Scaling up Standards-Based Accountability in the South 6
State Context and State Government 6
Study Purpose, Data Sources, Summary of Findings 9

Part I: State Politics and Accountability Policies in Virginia and Georgia Before NCLB 12
1995-2002: The Development of a Standards-Based Accountability System in Virginia 12
2001-2003: Virginia State Politics and Year 1 NLCB Implementation 15
1990s: Accountability Policies in Georgia from (D) Miller to (D) Barnes 18
2001-2002: Georgia State Politics and Year 1 NCLB Implementation 22

Summary of Comparisons Between Virginia and Georgia 24

Part II: Implications of Changes to State Accountability Systems 26
Question #1: Where did Virginia and Georgia define "proficiency"? 27
Question #2: What is a reasonable expectation for annual achievement gains, as measured by the NAEP basic level of performance? 33

Policy Recommendations 36

Discussion and Conclusion 37

List of Tables

Table 1: The Educational Governance Structure in Virginia and Georgia 7
Table 2: Provisional Accreditation Benchmarks for Virginia Middle Schools 12
Table 3: Number of Virginia Public Schools by Accreditation Level, 1998-2003 13
Table 4: 10-Year Gains (1990-2000) at NAEP Basic in Grade 8 Math by Ethnicity 38

List of Figures

Figure 1: Comparison of Grade 8 Math (2000) Proficiency Rates on the Virginia NAEP and Pass Rates ("Pass-Proficient") on the Virginia SOL 27
Figure 2: Comparison of Grade 8 Math (2000) Proficiency Rates on the Georgia NAEP and Pass Rates ("Meets State Expectations) on the Georgia CRCT 28
Figure 3: Comparison of Grade 8 Math (2000) Basic Rates on the Virginia NAEP and Pass Rates ("Pass-Proficient") on the Virginia SOL 29
Figure 4: Comparison of Grade 8 Math (2000) Basic Rates on the Georgia NAEP and Pass Rates ("Meets State Expectations") on the Georgia CRCT 30
Figure 5: Four-Year NAEP Trends (1992-1996, 1996-2000) and Four-Year SOL Trends (1998-2002) by Racial and Ethnic Subgroups in Virginia 33

Introduction: NCLB, State Responsibilities, and Civil Rights[1]

The "No Child Left Behind" Act (NCLB) of 2001 expands the federal role in K-12 education and enlarges the responsibility of state education agencies (SEA)[2] in implementing a standards-based accountability system. It requires states to develop challenging academic content standards that contain "rigorous content" and encourages the teaching of "advanced skills" (*Federal Register*, Vol. 67, No. 129, Section 200.1, 2002). Starting in 2005-2006, annual assessments in reading and math from grades 3 to 8 are required to assess student mastery of the standards. Performance standards must be applied to the assessments and include at least three achievement levels--basic, proficient, and advanced—describing how well students have learned the academic standards. The centerpiece of NCLB—and perhaps the most contentious requirement—is the strict 12-year timeline for ensuring that all students and subgroups defined by minority and low-income status reach the state-defined "proficiency" standard. To monitor progress toward this goal, NCLB requires states to establish adequate yearly progress (AYP) goals to ensure that all students and major demographic subgroups reach "proficiency" within 12 years (2013-2014). Schools failing to make AYP targets for any given subgroup for two consecutive years will be identified as "needing improvement" and, thus, subject to a series of sanctions, ranging from public school choice to school reconstitution.

Although NCLB permits states to define "proficient" performance on the state assessment, many states set their performance levels prior to the enactment of NLCB and were unaware of the consequential decisions that would hinge on the proficiency definition. State policymakers, for example, did not know that proficient levels set prior to NCLB would

---

[1] Special thanks to Mei Mei Peng, Khadijah Salaam, and Kate Sobel for providing excellent research assistance.
The author can be contacted at jimmykim@law.harvard.edu
[2] The terms "state education agency" and "state department of education" are used interchangeably throughout the paper.

determine the initial starting point for improving achievement over 12 years and the annual gains needed to bring all students and racial minorities to 100% proficiency. As a result, the climb toward the 100% proficiency target depends largely on the proportion of students meeting the proficient standard prior to 2002. Moreover, in the first year of NCLB implementation, states were also required to identify Title I schools that failed to meet definitions of AYP, which were determined before NCLB and varied dramatically across states. Schools that fail to meet AYP targets for each subgroup of students will be subject to a series of sanctions that become more punitive over time.[3] The stated goal of disaggregating achievement results is to hold schools responsible for improving learning outcomes for minority students and other disadvantaged subgroups.

Recent scholarship, however, has shown that the unintended consequences of the AYP requirements may adversely impact schools with larger minority enrollments. For example, analyses by Kane and Staiger (2002, p. 258) suggest that integrated schools will be disproportionately and unfairly sanctioned for failing to meet AYP targets. There are two reasons for this. First, such schools will be required to meet more than one AYP target and, second, scores for small subgroups of students are volatile and yield unreliable measures of school performance. In many ways, then, racial equity underlies NCLB's stringent achievement targets and accountability mandates. Therefore, it is important for researchers to examine the civil rights implications of NLCB's assessment and accountability requirements, especially in states with large minority enrollments.

---

[3] P.L. 107-110, Section 1116 (b)(5-9). If a school fails to meet AYP for two consecutive years, it must develop a "school improvement plan" and students must be given the choice to attend another school in the district. Students in schools failing to meet AYP for three consecutive years may also use Title I dollars for supplemental services, such as private tutoring. After three years of failure, schools are subject to "corrective action" and "restructuring," which may result in one of several sanctions including the replacement of the school staff and state take over of the school.

Study Purpose, Data Sources, Summary of Findings

This study has two broad goals—to discuss state politics and accountability policies in Virginia and Georgia before NCLB, and to analyze the civil rights implications of NCLB's accountability requirements for future policy decisions. In part one, we describe previous policies and governance arrangements and how these factors affect state capacity to meet NCLB's accountability requirements. We rely on state policy documents, newspaper articles, and interviews with state officials to provide a short narrative history of Virginia and Georgia's efforts to construct an accountability system during the 1990s. Before NCLB, Virginia had developed a coherent set of accountability policies, and its newly elected governor, the State Board of Education, and the State Education Agency worked collaboratively to implement major accountability requirements under NCLB. Since the state's standards and assessments met all the 1994 Title I requirements, state policymakers focused attention on plans for improving reading instruction statewide and building the instructional capacity of the small number of low-performing schools identified for improvement under NCLB. Unlike Virginia, Georgia experienced political conflicts between the Governor and Superintendent of Instruction. This delayed the state's progress toward building a coherent accountability system and developing a coordinated strategy for implementing NCLB. By 2002, the state had to revamp its curriculum, construct new high school tests, rebuild the capacity of the Georgia Department of Education to implement the new Title I requirements, and restore working relations among leaders of various state education agencies.

Part two of the study analyzes the implications of NCLB's accountability requirements for state policy decisions, especially with respect to establishing AYP starting points and improvement goals. In response to NCLB, both states adopted the 12-year timeline for having

all students become proficient on the state assessment. As a result, both Virginia and Georgia set initial pass rates and adequate yearly progress (AYP) targets for bringing all students to the proficient level. We use achievement results from the state test and NAEP to compare Virginia and Georgia's proficient levels on the state assessment with the NAEP performance standards in grade 8 math. We find that proficient performance on the Virginia and Georgia state test corresponds to the "basic" level of performance on NAEP for all major subgroups. Although it may be reasonable for states to improve the percentage of students meeting the NAEP basic level on an annual basis, we highlight several assumptions that would have to be met in order to reach these goals. We also discuss the civil rights implications of the racial subgroup rules, which may unfairly disadvantage racially diverse school.

NCLB landed on states with varying degrees of alignment with the new Title I accountability mandates. The initial response of two states—Virginia and Georgia—suggests that implementation of NCLB's prescriptive accountability requirements will depend on the (1) cohesiveness of state accountability policies before the enactment of NCLB and (2) coordinated policymaking between the State Board of Education (SBE), the State Education Agency (SEA), and state political leaders.

Part I: State Politics and Accountability Policies in Virginia and Georgia Before NCLB

State Context and State Government

The experiences of two states—Virginia and Georgia—provide in-depth case studies of initial state efforts to implement NCLB. Virginia has made more progress than Georgia in developing an accountability system that already incorporates the requirements of the federal law. For example, Virginia was one of only 19 states that fully complied with the 1994 assessment requirements (Robelen, 2002). In 1995, the State Board of Education adopted Standards of Learning (SOL), and phased in assessments aligned with the new state curriculum. The SOL assessments underpin the state accountability system (Goertz, Duffy, & Le Floch, 2001), and the Title I accountability system is already aligned with the state's main accountability system. According to Ravitch (2002), "Virginia has been a national leader in the development of rigorous standards and tests and on setting accountability benchmarks for both students and schools" (p. 12). Georgia, by contrast, is farther behind in its starting point relative to what the new law requires. The state received a waiver to comply with the 1994 Title I accountability requirements. To be in compliance, the state must replace the minimum competency graduation exams with end-of-course tests (Cohen, 2001). In short, Georgia has made substantially less progress than Virginia in developing a coherent testing and accountability system that satisfies the Title I requirements of NCLB. According to one education policymaker, Georgia has "not done a lot of things that other Southern states have done to develop a standards-driven system. They're playing catch-up" (Jacobson, 2003).

Virginia and Georgia offer "critical case" studies that have implications for the other 50 states. According to Patton (1990), "the clue to the existence of a critical case is a statement to the effect, 'if it happens there, it will happen anywhere,' or, vice versa, 'if it doesn't happen

there, it won't happen anywhere'" (p. 174). Georgia embodies the first statement. Given the instability of the political environment and the lack of policy coordination in Georgia, it represents a state where policymakers have encountered obstacles in developing a coherent accountability system. Virginia, on the other hand, is a state where policymakers already incorporated many of the underlying principles in NCLB into the state accountability system. To a large extent, Virginia and Georgia face different implementation challenges related to the constitutional constraints placed on state policymakers.

The distribution of power is more centralized in Virginia than Georgia, facilitating the development of coherent education policies. For example, in quantifying the degree of state centralization in education policymaking, Wirt (1977) found that power was more centralized at the state level in Virginia than in Georgia. Additional contrasts between the governance structure in Virginia and Georgia are illustrated in Table 1 below. Virginia's Constitution

Table 1: Educational Governance Structure in Virginia and Georgia.

	Virginia	Georgia
State Board of Education	appointed by Governor 9 members	appointed by Governor 13 members, one for each congressional district
	4-year terms serve staggered terms	7-year terms served staggered terms
State Superintendent	appointed by Governor "experienced educator"	elected official chief executive, Board of Education

Source: Virginia and Georgia State Constitution.

(Article VIII, Section 4) authorizes the Governor to appoint the Superintendent of Public Instruction and all 9-members of the State Board of Education. Since each board member serves four year, staggered terms, this potentially allows the governor to appoint a majority of any given board. The Superintendent of Public Instruction must be an "experienced educator" who is

responsible for governing the administrative functions of the State Department of Education (Virginia Constitution, Article VIII, Section 6).

The Georgia Constitution also authorizes the governor to appoint members to the Board of Education, which includes 13 members, one for each congressional district, giving equal representation to all local jurisdictions. Since Board members serve for 7 years and since there are 13 members, it is difficult for any single Governor to appoint a majority of the Georgia Board of Education (Georgia Constitution, Article VIII, Section 2). The Georgia Constitution (Article VIII, Section 3) also requires that the State Superintendent of Instruction be "elected at the same time and in the same manner and for the same term" as the Governor. The State School Superintendent is the executive officer of the State Board of Education. In principle, the Georgia Constitution builds in a strong system of checks and balances by creating one Board member for each congressional district and by authorizing the election of a State Superintendent of Instruction. As a result, the Governor is less likely to have a solid majority on the State Board and cannot choose a State Superintendent who shares his political views and supports his policy agenda. Compared to Virginia, authority to make education policy in Georgia is diffused among the Governor's office, the State Board of Education, and the elected State Superintendent of Instruction. If political conflicts arise, however, the Governor may ultimately avoid working with the Board and Superintendent by initiating education legislation in the General Assembly. Nevertheless, effective implementation of newly enacted education laws ultimately requires collaboration across the three major policymaking entities—the Governor's office, the State Board of Education, and the State Education Agency.

1995 to 2002: The Development of a Standards-Based Accountability System in Virginia

Virginia began to develop a coherent accountability system in the mid-1990s by revising its academic content standards, adopting criterion-referenced tests tied to the standards, and linking school accreditation and high school graduation to performance on the statewide assessments. In 1995, the Board of Education adopted new K-12 Standards of Learning (SOL) in reading, math, science, history/social science. Criterion-referenced tests linked to the SOLs were administered in the spring of 1998, and cut scores defining three achievement levels—basic, proficient, and advanced—were applied to the SOL scores and end-of-course high school exams in the fall of 1998. High-stakes were tied to SOL performance for both individual students and schools. However, the Board chose to phase in the sanctions over time. For example, by 2003-2004, students had to pass six end-of-course high school exams to graduate, and by 2006-2007, schools had to earn 70% pass rates in English, math, science, history/social science to be accredited. The school accreditation policy, therefore, would take effect 12-years after the enactment of the SOLs in 1995, allowing one cohort of public school students to be exposed to the new learning standards. Both the 12-year timeline for school improvement and the use of an absolute performance standard applied to all schools anticipated the accountability model underlying NLCB. Although few schools met accreditation standards during the first two administrations of the SOLs (Portner, 1999), the State Board of Education responded to political pressures from low-scoring schools and districts by modifying rather than dismantling the accountability system.

In 2000, the Board adopted several changes to the standards of accreditation, and, in retrospect, these revisions to state policy incorporated many of the principles and ideas embedded in NCLB. Echoing the language of "adequate yearly progress," the Board set annual goals that gradually increased in a "stairstep" pattern from 2000-01 to 2003-04. Table 2 below

highlights the Board's changes to Virginia's school accreditation policies, which established provisional accreditation benchmarks for middle schools (Virginia Board of Education, 2000).

Table 2: Provisional Accreditation Benchmarks for Virginia Middle Schools.

	1999-00	2000-01	2001-02	2002-03
English	60%	63%	66%	70%
Math	55%	60%	65%	70%
Science	60%	63%	66%	70%
History/Soc. Studies	40%	45%	50%	55%

Source: Virginia Standards for Accrediting Public Schools (2000).

These increases depended on the baseline SOL pass rates from 1999-2000. For middle schools, the baseline pass rates were 60% in English, 55% in math, 60% in science, and 40% in history/social science, and the percentage of students passing each SOL subject test had to increase in subsequent years. Although the expected improvements in pass rates varied by subject, the underlying principles of setting a baseline, intermediate goals, and a timeline for improvement were similar to NCLB's requirements for schools to show adequate yearly progress from an initial starting point in 2002-03 to 100% proficient in 2013-2014. To improve the reliability of school pass rates, the revisions allowed ratings to be determined by the higher of two numbers: three-year rolling averages or the current year's schools. Such an option also anticipated NCLB's allowance for the use of three-year averages.

The addition of annual targets and the provisionally accredited designation created incentives for schools to focus attention on improvements over time and allowed for more fine-grained distinctions among schools. Prior to 2000, a school that missed the 70% cut scores in one of four tests or by a small or large margin was essentially treated as "failing to meet accreditation standards." By creating four performance levels to categorize schools, the Board could focus resources on schools in the lowest performance category, "accredited with warning"—that is, schools with pass rates generally below 50% on all four SOL subjects. In complying with the 1994 IASA's requirement for states to define "adequate yearly progress" for

Title I schools, Virginia defined schools as failing to make AYP if they were labeled "accredited with warning." Title I schools failed to meet AYP if they remained in the "accredited with warning" category in English and/or math for two consecutive years (Consortium for Policy Research in Education, 2000a). By 2002, when the stricter NCLB requirements took effect, Virginia had identified a relatively small number of schools (34) for improvement. Most schools had shown improvement in SOL scores since 1998, leading to large increases in the number of schools meeting "provisional accreditation" benchmarks and large reductions in the number of schools "accredited with warning." Table 3 shows how the increase in accredited schools coincided with a drop in low-performing schools. By modifying the SOL-based accountability system over time, Virginia largely kept intact the core of its accountability system. And as schools became more familiar with the SOLs and the assessments, scores increased. By 2002, only 34 schools were identified for improvement, and the state was able to invest resources in the most disadvantaged schools.

Table 3: Number of Virginia Public Schools by Accreditation Level, 1998-2003.

Year	Total Schools	Fully Accredited	%	Provisonally Accredited-Meets State Standards	%	Provisionally Accredited-Needs Improvement	%	Accredited With Warning	%
1998-99	1772	39	2.2%						
1999-00	1787	117	6.5%	926	51.8%	N/A	N/A	697	39.0%
2000-01	1824	415	22.8%	713	39.1%	459	25.2%	174	9.5%
2001-02	1839	735	40.0%	557	30.3%	398	21.6%	95	5.2%
2002-03	1829	1180	64.5%	254	13.9%	310	16.9%	85	4.6%

Source: Virginia Department of Education.

The state undertook several approaches to helping low performing schools. For example, schools with pass rates below 50% on all SOL tests ("accredited with warning") were required to undergo an "Academic Review" by a team of state department officials. The purpose of the review was twofold: (1) to diagnose weaknesses in curriculum, instruction, and the allocation of school staff and time, and (2) to prescribe a plan for school improvement to encourage school

progress toward meeting the state accreditation requirements by 2006-2007. Moreover, the lowest-performing schools were also required to adopt instructional programs in reading and math that had a documented track record of success. Similarly, under NCLB, schools failing to make AYP must adopt interventions that incorporate "scientifically-based research," a term mentioned over 100 times in the federal law (Center on Education Policy, 2003).

2001-2003: Virginia State Politics and Year 1 NLCB Implementation

Before the enactment of NLCB, Virginia's school accountability was substantially aligned with the federal Title I requirements, and state leaders—in particular, the Governor, the President of the Board of Education, the State Superintendent of Instruction—had a coordinated agenda for complying with NCLB. The overriding objective of state leaders was to build the instructional capacity of "schools accredited with warning." For example, in fall 2002, the newly elected governor, (D) Mark Warner, continued the SOL-based reforms of his Republican predecessors by proposing a $3 million education initiative for improving SOL pass rates in the 34 schools identified for improvement under NCLB (O'Dell, 2002). Warner's Partnership for Achieving School Success (PASS) program sends educators and volunteers from business and civic organizations to low-performing schools that need assistance in improving SOL scores. PASS enhances the state accountability program by supporting the capacity of low-performing schools to upgrade curriculum and instruction. The first pairing between low-performing schools and corporate sponsors and volunteers took place in the 34 Title I schools that were identified by NCLB for failing to make AYP for two consecutive years. PASS, therefore, has a fairly narrow policy focus since it targets Virginia's lowest-performing schools.

In addition to sustained support for the accountability system among Republican and Democratic administrations, the current Board of Education President, Mark Christie, has been

involved in the design, implementation, and modification of the accountability system since its inception in 1995. Citing flat SOL pass rates in English and concern over Virginia's performance on the NAEP reading, Christie led recent Board efforts to create "A Committee to Implement NCLB" that focused on improving reading instruction. The first meeting brought in Reid Lyon, The Director of the National Institute of Child Health and Human Development, who shared the findings of the National Reading Panel (Virginia Board of Education, 2002). These findings, which emphasized the need for teachers to focus instruction on phonemic awareness, phonics, fluency, vocabulary, and comprehension strategies, provided a blueprint for improving reading instruction statewide (Virginia Board of Education, 2002). Subsequently, throughout fall 2002, the Board focused on teacher training as a primary strategy for scaling up improved reading instruction statewide.

State implementation of NCLB was also facilitated by Governor Warner's re-appointment of the State Superintendent of Instruction, Jo Lynne DeMary, who was originally appointed by Republican James Gilmore. According to Warner, continuity in leadership was needed "at a time when new testing and accountability mandates are coming from Washington, Dr. DeMary will provide stable leadership and keep Virginia ahead of the curve on education reform" (Associated Press, 2002). The selection also won statewide support because of DeMary's "apolitical" approach to education policy. One editorial in the Virginian-Pilot observed that "while some superintendents have used the post as an ideological pulpit or a conduit to more lofty ambitions, DeMary approaches the job as the capping of a career that has focused first and foremost on the Virginia school system" (The Virginian-Pilot, 2002). From a policy perspective, DeMary's re-appointment provided stable leadership in Virginia's Department of Education, which had to work on the nuts and bolts details of NCLB's

accountability provisions, such as the development of initial starting points on the 12-year timeline. Indeed, the completion of the June 20, 2002 Consolidation Application for ESEA grants reflected collaboration among various divisions within the Virginia Department of Education (Virginia Department of Education, 2002).

1990s: Accountability Policies in Georgia from (D) Miller to (D) Barnes

Throughout much of the 1990s, the development of academic content standards in Georgia was largely divorced from assessment policy. Standards-based reforms originated in 1986, when the Georgia General Assembly passed legislation creating a statewide curriculum. Called the Quality Core Curriculum (QCC), it is the counterpart to the Virginia SOLs. The legislation required districts to administer norm-referenced tests in grades 2, 4, 7, and 9. The tests, however, were not aligned with the QCC. As a result, Georgia's definition of "adequate yearly progress," as required by Title I of the 1994 Improving America's School Act (IASA), was based on a norm-referenced test rather than a criterion-referenced test linked to the Georgia QCC (Consortium for Policy Research in Education, 2000b). Furthermore, Georgia's high school assessment—a minimum competency graduation test—did not assess student mastery of the high school QCC. Although Georgia planned to revamp its high school exit exam and replace it with more rigorous end-of-course tests in English, math, science, and social studies, the state failed to implement the new assessments by 2001. Thus, to complete administration of the new high school assessments, Georgia received a 2-year federal waiver. Under the waiver, Georgia had until 2003 to build a final assessment system that met the requirements of the 1994 IASA (Cohen, 2001). In some ways, the inability of the state to build a coherent K-12 accountability system reflected the priorities of Georgia's governor, (D) Zell Miller, who, throughout much of the 1990s, focused on improving access to pre-K programs and expanding the Hope scholarships for college students. Consequently, Miller's administration placed less emphasis on developing K-12 accountability policies (Watts, 2003).

Georgia first began to develop an accountability system in 1997 when the State Board of Education led a major effort to revise the QCC and awarded a contract to Riverside Publishing to

develop criterion-referenced competency tests (CRCT). The goal was a test that measured student mastery of the QCC in grades 4, 6, and 8. Then, in 2000, Governor Barnes proposed legislation that would hold students and schools accountable for meeting performance standards applied to the CRCT.

The Governor's plan, called the A-Plus Education Reform Act (HB 1187), was passed by the Georgia legislature in July 2000. This bill contained three key elements that would establish the constraints for implementing NCLB. First, the bill authorized the governor to create a new Office of Educational Accountability (OEA), which would be controlled by the governor's office rather than the Georgia State Department of Education. Second, the new law gave OEA the authority to create an "A" to "F" school grading system that would be based on CRCT scores. Third, the law expanded CRCT testing in reading and math in grades 1 to 8, and in science and social studies from grades 3 to 8. The law also phased out the high school graduation test and replaced them with end-of-course tests in grades 9-12. The first administration of the CRCT took place in grades 4, 6, and 8 during spring 2000. Based on spring CRCT scores, standard setting panels created three performance levels that correspond to basic (does not meet state expectations), proficiency (meets state expectations), and advanced (exceeds state expectations). One year after the enactment of HB 1187, the Georgia legislature also passed Governor Barnes' proposal to end social promotion. Phased in over time, the new policy set promotion criteria based on the CRCT in grade 3 reading (effective spring 2004), grade 5 reading and math (effective spring 2005), and grade 8 reading and math (effective spring 2006).

HB 1187 represented the most ambitious education reform bill in Georgia's history. It required collaboration between the Governor's Office of Educational Accountability, which had to design the school grading system, and the State Department of Education, which had to

develop the CRCTs and provide schools with resources to prepare for the CRCT. However, the Governor, (D) Roy Barnes, and the elected State Superintendent of Instruction, (R) Linda Schrenko, disagreed about strategies for improving Georgia's schools. Indeed, Schrenko had criticized Barnes' education reform bill, HB 1187, using the SDE's website to campaign against the governor's plan (Laccetti, 2000). By 2001, the increasingly adversarial relationship between Barnes and Schrenko undermined the collaborative planning that was needed to create a coherent accountability system. Given HB 1187's heavy emphasis on increased testing and accountability for individual schools and students, officials from Schrenko's State Department of Education notified the Board of Education in May 2001 that "the enormity of the task well exceeds the capability and capacity of the current Georgia Department of Education" (Phi Delta Kappa International, 2002). Concerned about the new CRCT-based accountability system, SDE officials argued that the state needed an approach that would "improve student performance of low performing schools quickly." It chose the National Center on Education and the Economy (NCEE) to implement a statewide comprehensive school reform model called America's Choice. However, America's Choice contained curricular areas that were not aligned with the Georgia's curriculum.

By expanding testing requirements, the architects of HB 1187 assumed that the curriculum was rigorous and that teachers had the resources and training to teach students for the high-stakes CRCT exam. Two independent reports, however, questioned these assumptions. First, the American Federation of Teachers (2001) pointed out that Georgia had not developed instructional resources in English, math, science, and social studies, and had no instructional strategies for three subjects (math, science, social studies). Second, Phi Delta Kappa International (2002), in a comprehensive audit of the QCC, concluded that "the QCC standards

fail to give adequate specificity for teachers or curriculum designers to clearly understand what students are to know to meet any given standard" (Phi Delta Kappa International, 2002). Furthermore, the CRCT assessments "must be more focused, some of which are not aligned or which are not universally administered across the state should be dropped" (Phi Delta Kappa International, 2002)

With respect to school accountability, the Title I AYP measure required schools to reduce the proportion of students at the "does not meet standard" level on the CRCT by 5-percentage points. This AYP definition, however, was not based on the school grading system, and it did not focus on increasing the percentage of students at the proficient level, as required by NCLB. In choosing a 5% move out rate from the basic level, Superintendent Schrenko and staff in the testing and evaluation division expected teachers to move at least one student out of basic level each year. Thus, schools with a large fraction of students meeting state standards could be identified for failing to make AYP if they did not reduce the proportion of students in the "did not meet standard" category by at least 5-percentage points. In 2002, when Georgia had to identify Title I schools that failed to make AYP, the CRCT-based accountability system had been in place for only two years. Consequently, in the summer 2002, 437 Georgia schools were identified for failing to make AYP and were required to offer choice and/or supplemental services, forcing the SDE to provide support in nearly 20% of all Georgia public schools.

2001-2002: Georgia State Politics and Year 1 NCLB Implementation

In the first year of NCLB implementation, deepening political conflict between Governor Barnes and Superintendent Schrenko overshadowed Georgia's substantive plans to implement NCLB. A major controversy arose in June 2002, when Superintendent Schrenko defied the wishes of the governor and State Board of Education by writing a letter to U.S. Secretary Paige to request that the U.S. Department of Education not consider Georgia for the State Flexibility Demonstration Program. Participation in this program would give the state increased flexibility in the use of federal dollars. Board Chair, Cathy Henson, responded incredulously that, "the superintendent has always said that she favors local control and flexibility in the use of federal and state funds, so it was a real shock to learn that she now favors federal control over how Georgia uses federal dollars" (Georgia Board of Education, 2002). By unanimous vote, the Board followed by passing a motion to override Shrenko's decision by sending a letter to Secretary Paige asking that Georgia be considered for the State-Flex program.

In addition to the conflict among Georgia's elected officials, the fragmentation of education policymaking exacerbated efforts to implement a coordinated strategy for complying with NCLB. For example, Barnes' major education bill, HB 1187, removed a number of administrative powers from the State Department of Education, including control over accountability policies, teacher certification, and data analysis. As a result, 10 different state agencies were involved in drafting Georgia's June 2002 "Consolidated Application" for ESEA funds. The responsibility for developing specific accountability policies, such as defining adequate yearly progress, rested with three difference offices--the Office of Educational Accountability, the Education Coordinating Council, and the State Board of Education. By fall 2002, Georgians voted into office a new Governor, (R) Sonny Perdue, and new State

Superintendent of Instruction, (R) Kathy Cox. Both vowed to work together in rebuilding the Georgia SDE and restoring collaboration among the difference state agencies responsible for education policy (Jacobson, 2003).

From the perspective of Georgia's elected officials, making coherent state policies that facilitated implementation of NCLB first required the state to dismantle the educational bureaucracy created by former Governor Barnes. This re-organization began in earnest on February 18, 2003, when Governor Perdue and Superintendent Cox proposed an education bill called STARS (Students + Teachers + Accountability + Respect = Success), which authorized the Governor to return the Office of Education Accountability and the Student Data and Research Center to the Department of Education. Such a plan, according to the new Superintendent Cox, would make the Department a "one-stop agency for education" and "eliminate the duplicate bureaucracy that has stripped control from those at the local level and left them confused as to who was in control of certain areas" (Georgia Department of Education, 2003). In addition to returning control over the implementation of education policies to one agency, Governor Perdue moved quickly to appoint new members to the Board of Education, including a new Chair of the State Board of Education (Salzer, 2003). With the appointment of Chair Wanda Barrs, the Governor had a new team in place. The sweeping changes in Georgia's education establishment promised to usher in, what Superintendent Cox, called a "new beginning for Georgia's schools," adding that "the era of divisiveness is over. The Governor, the Board of Education, and the Department of Education are all united for the purpose of raising student achievement across the state" (Georgia Department of Education, 2003). Political alignment, Cox believed, was a prerequisite for creating a coordinated agenda for improving Georgia's

schools. Thus, in the first year following NCLB, Georgia's new political leaders moved quickly to re-organize the state agencies in charge of education policy.

Summary of Comparisons Between Virginia and Georgia

The disparity in the number of schools identified for improvement in Virginia and Georgia reflect much deeper differences in each state's accountability policies and political environment. Virginia has been able to focus efforts on improving the instructional capacity of the state's lowest achieving schools. The state's "Committee to Implement NCLB" focused attention on improving reading instruction, the SEA's school improvement teams provide concrete plans for helping schools meet accreditation standards, and the Governor's first major education bill (PASS) largely reinforced these efforts. The 34 Title I schools identified for improvement represented fewer than 2% of all Virginia public schools so the state is able to concentrate money, time, and resources in the most disadvantaged schools. With the exception of Richmond City Schools, which had the most schools identified for improvement, very few schools and districts had to offer choice and supplemental services. There is, in short, a strong link between state policy and local efforts to improve teaching and learning at the school level.

Prior to NCLB, Georgia made less progress in developing a coherent accountability system. Deep political divisions between the former Governor and his appointees to the Board of Education, on one hand, and the former State Superintendent of Instruction and the SEA, on the other constrained state action on accountability. The political conflict between these two elected officials created fragmented policies focused more heavily on increasing testing and accountability than on improving the curriculum and instructional supports for teachers. The major education bill introduced under the Barnes administration further fragmented the administrative and policymaking structure in the state. Thus, the newly elected governor's first

major education bill, STARS, focused more on rebuilding the state's education bureaucracy than on programs for improving curriculum, instruction, and achievement. Since nearly one-fifth of all Georgia public schools were identified for improvement, it may be increasingly difficult for the SEA to provide technical assistance to all 437 Title I schools. Efforts to build instructional capacity in these schools may be hampered as educators implement choice and supplemental services—interventions that may fragment school-wide efforts to build instructional coherence. If the goal of NCLB is to focus resources on the lowest-performing schools, Virginia has made more progress toward this goal than Georgia. Despite this key difference, NCLB ultimately required both states to set a 12-year timeline for ensuring that all students reached proficiency. The implications of this requirement for state policy are discussed in part II.

Part II: Implications of NCLB's Accountability Requirements for State Policy

Since the enactment of NCLB, both Virginia and Georgia responded by keeping intact pre-existing policies that met current federal testing and accountability requirements. For example, both states have statewide assessments aligned to academic content standards. Each state also has three performance levels on its state test, including a "proficient" cut score, which corresponds to the "pass-proficient" label on the Virginia SOL and the "meets state expectations" label on the Georgia CRCT. Virginia has not changed its definition of proficiency since it was first applied to the first administration of the SOL tests in spring 1998. Georgia's definition of proficiency on the CRCT has remained the same since spring 2000. However, both Virginia and Georgia defined proficiency levels on their statewide assessments long before the enactment of NCLB. How high or low each state defined proficient performance will determine the initial starting point and the magnitude of gains needed to meet the 100% proficient target.

Part II of this paper addresses two questions prompted by the 12-year requirement for ensuring that all students reach proficiency on the state assessment. First, how does the state definition of proficiency compare to the NAEP performance levels? To address the first question, we compared the percentage of students performing at proficiency on the SOLs and CRCTs to the NAEP basic and proficiency level in grade 8 math, since both Virginia and Georgia administered statewide assessments and participated in NAEP in grade 8 math. Although NCLB does not require states to make annual improvements on NAEP, it does encourage states to verify state gains with trends on NAEP. Indeed, we found strong correspondence in the percentage of students meeting the state-defined proficiency level and the basic level on NAEP. Thus, it may be reasonable to expect states to improve not only on the state assessment but also on NAEP. Hence the second question we address is: what is a

reasonable expectation for annual achievement gains, as measured by the NAEP basic level of performance?

Question #1-Where did Virginia and Georgia define "proficiency"?

Using grade 8 math achievement data from spring 2000, we compared the definition of proficiency on the Virginia SOLs and Georgia CRCTs to results from NAEP. We focused on grade 8 math results from spring 2000 for two reasons. First, both Virginia and Georgia administered state assessments in eighth-grade math and participated in the eight-grade NAEP in 2000, allowing us to compare state performance standards to an external standard. NAEP results, therefore, provide a common yardstick for judging state performance standards. Second, by 2000, both Virginia and Georgia had established three achievement levels, which correspond to the basic, proficient, and advanced levels of performance on NAEP.[4] The following results show the proportion of students meeting the state's proficiency standard with the proportion meeting the (1) NAEP proficient and (2) NAEP basic standard.

Figure 1 displays the proportion of students meeting the proficiency standard on NAEP and the Virginia SOLs. Starting from the left of Figure 1, the first two bars shows that 26% of eighth-graders reached the NAEP proficiency standard, but nearly three times as many eighth-graders (61%) met the proficient standard on the SOLs. The next set of results compares NAEP and SOL results for each of the four ethnic groups. The difference in the height of each pair of bar graphs clearly shows a higher proportion of students meeting the proficiency standard on the SOLs than NAEP. Although the disparity in state and NAEP results applies to all ethnic groups, it is substantially larger for Black and Latino students than for White and Asian students. For

[4] In particular, the "pass-proficient" standard in Virginia and the "meets state expectations" standard in Georgia are nominally related the proficient standard on NAEP. Neither Virginia nor Georgia has changed the criterion needed to pass the state test since 2000. In Virginia, students had to answer 37 out of 60 (62%) to pass the eight-grade SOL math.

example, nearly seven times as many Blacks reached the state proficient level (36%) compared to the NAEP proficient level (5%), and about four times as many Latino students met the state standard (53%) as compared to the NAEP standard (14%). The discrepancy between state and NAEP performance in general and the substantially larger discrepancy among Black and Latino students is also shown in the results for Georgia.

Figure 1: Comparison of Grade 8 Math (2000) Proficiency Rates on the Virginia NAEP and Pass Rates ("Pass-Proficient") on the Virginia SOL.

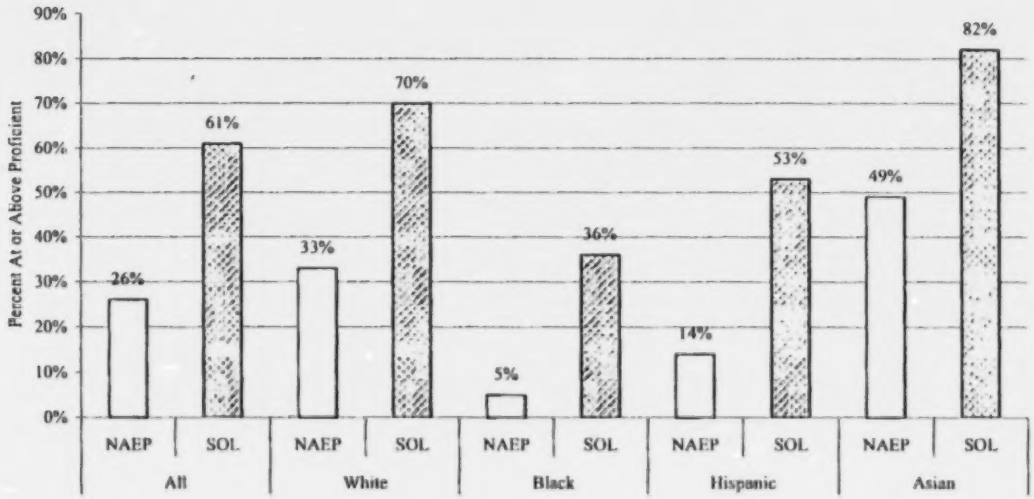
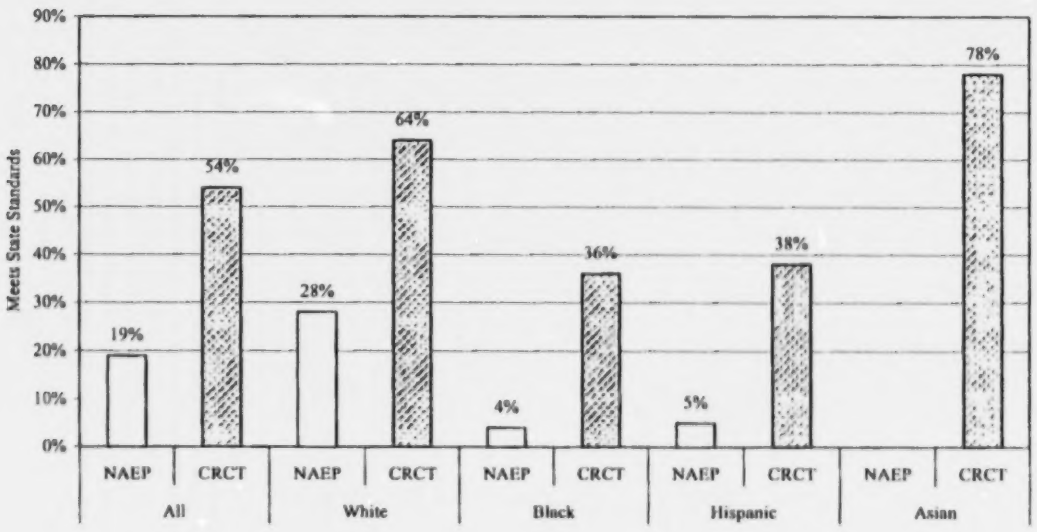


Figure 2 shows the percentage of eighth-graders in Georgia who scored at or above proficient on the CRCT and NAEP. The first two bars show that 19% of eighth-graders met the proficient standard on NAEP compared to 54% on the Georgia CRCT. Divergence in the proportion of students reaching NAEP and CRCT proficiency level also appears when the results are broken down by race and ethnicity. For Whites, proficiency rates are about twice as large on the CRCT (64%) than NAEP (28%). Furthermore, proficiency rates are nearly nine times larger on the CRCT than NAEP for Blacks (36% vs. 4%) and eight times larger for Latinos (38% vs.

5%). Although the specific proportion of students meeting proficient levels varies in Georgia and Virginia, the general pattern of score trends is quite similar. In sum, substantially larger fractions of students meet proficiency as defined by the state than NAEP, and the disparity between results on the state test and NAEP is larger for Black and Latino students than for White and Asian students.

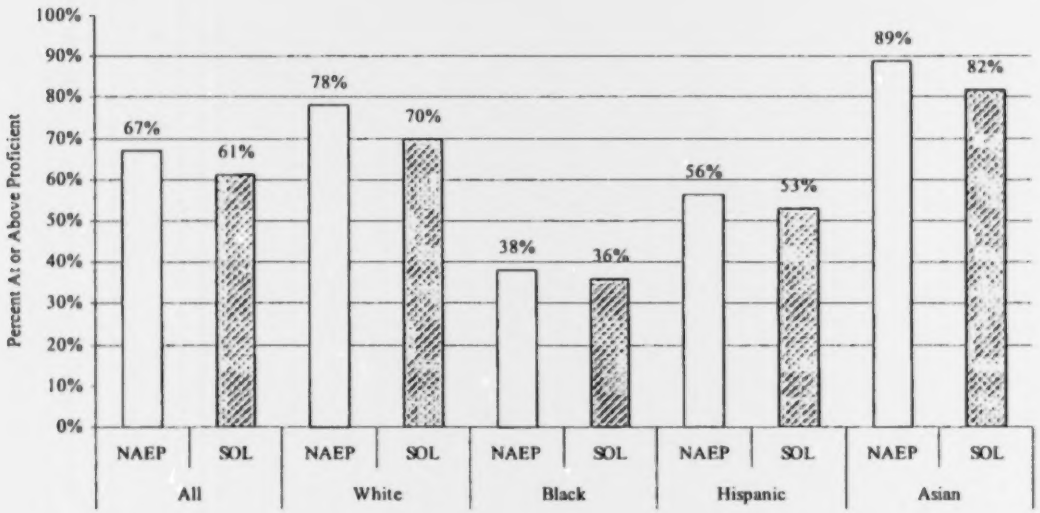
Figure 2: Comparison of Grade 8 Math (2000) Proficiency Rates on the Georgia NAEP and Pass Rates ("Meets State Expectations") on the Georgia CRCT



Instead of using the NAEP proficient standard as a criterion for judging the level of state performance standards, the next two figures compare state proficiency standards to the NAEP basic. Figure 3 shows the percentage of students meeting the NAEP basic standard and the SOL proficiency standard. As shown by the similar height of each pair of bars, we see near exact agreement between the proportion of students meeting the NAEP basic standard and the state proficiency standard. Broken down by ethnicity, the proportion of students reaching proficiency

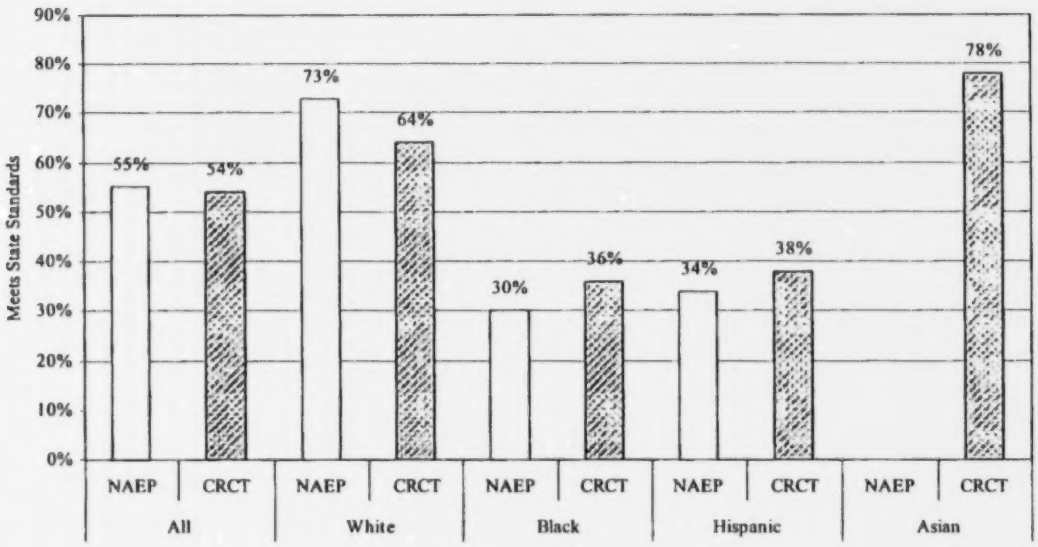
level on the Virginia SOLs and the NAEP basic standard varies only by 7% for Whites (70% vs. 78%), 2% for Blacks (36% vs. 38%), 3% for Latinos (53% vs. 56%), and 7% for Asians (82% to 89%).

Figure 3: Comparison of Grade 8 Math (2000) Basic Rates on the Virginia NAEP and Pass Rates ("Pass-Proficient") on the Virginia SOL



Substantial agreement in the proportion of students meeting state proficient levels and NAEP basic also apply in Georgia. Figure 4 juxtaposes the percentage of students scoring at or above basic on NAEP with the percentage of students who pass the Georgia CRCT. The first two bars show nearly identical proportions of students who met the NAEP basic standard and the Georgia proficient standard. Similarly, for White students, there is only a small difference between the percentage of students meeting either standard. For minority students, a little over 30% of all eighth-graders met the NAEP basic standard and the "meets expectations" threshold on the CRCT.

Figure 4: Comparison of Grade 8 Math (2000) Basic Rates on the Georgia NAEP and Pass Rates ("Meets State Expectations") on the Georgia CRCT



Although NCLB requires all students to reach "proficiency" by 2014, NAEP's definition of proficiency is substantially higher than the corresponding proficient label attached to either the Virginia SOLs or Georgia CRCTs. In many respects, the divergence is inevitable given the judgmental process that led to the creation of NAEP's performance levels, which represent unrealistically high performance levels that few students are able to attain (Linn, Koretz, & Baker, 1996). The NAEP performance standards were set so high that nearly one-fourth of students in high-achieving countries like South Korea and Japan would fail to meet the standard (Linn, 2000). In judging proficiency standards set by Virginia and Georgia, it appears that both states unintentionally set proficiency levels that correspond to the NAEP basic level for all students as well as for White, Black, Latino, and Asian students.

However, in both Virginia and Georgia, substantial correspondence between the percentage of students meeting state "proficiency" levels and the NAEP "basic" is probably incidental rather than intentional. Both states relied on judgmental standards-setting processes that involve the subjective opinions of numerous stakeholders (e.g., teachers, administrators, state policymakers), each of whom had quite different notions of what proficient performance entails and the number of questions a proficient student should answer correctly on a standardized test. Yet the striking similarities in the percentage of student meeting the state-defined proficiency standard and the NAEP basic level, as shown in Figure 3 and 4 suggest that, at minimum, state performance standards should be as rigorous as the NAEP basic level of performance. Since NCLB requires all states to move 100% of students to the proficient level on a state test, strong incentives exist for states to comply with NCLB by lowering proficiency standards and creating a dual system of accountability—that is, a set of proficiency standards for state accountability and another for federal accountability. For example, Colorado will count students reaching the "partially proficient" standard on the CSAP as meeting the federal definition of "proficient." Connecticut recently established a new proficiency level to comply with federal law, but it will be lower than the state's definition of proficiency (Education Week, 2002).

Inevitably, more states will adjust their performance standards to establish more realistic—and perhaps easier—targets for having all students meeting proficiency within 12 years. Conversely, states with lower proficiency definitions may have little incentive to raise the bar. NCLB requires states receiving Title I funding to participate in biennial administrations of the NAEP in 4th and 8th grade reading and math. In August, the Secretary of Education proposed a new regulation mandating that local districts receiving Title I funds participate in NAEP,

replacing a prior law allowing for voluntary district participation (*Federal Register*, Vol. 67, No. 151, Section 200.11, 2002). Although NCLB does not require states to show improvement on NAEP, federal officials, researchers, policymakers, and journalists will inevitably use NAEP as a national yardstick for judging state standards. According to one measurement specialist, NAEP is needed because "it will be almost impossible to get a national picture from the state assessments. This is where NAEP comes in. That's what people will rely on" (Hoff, 2002).

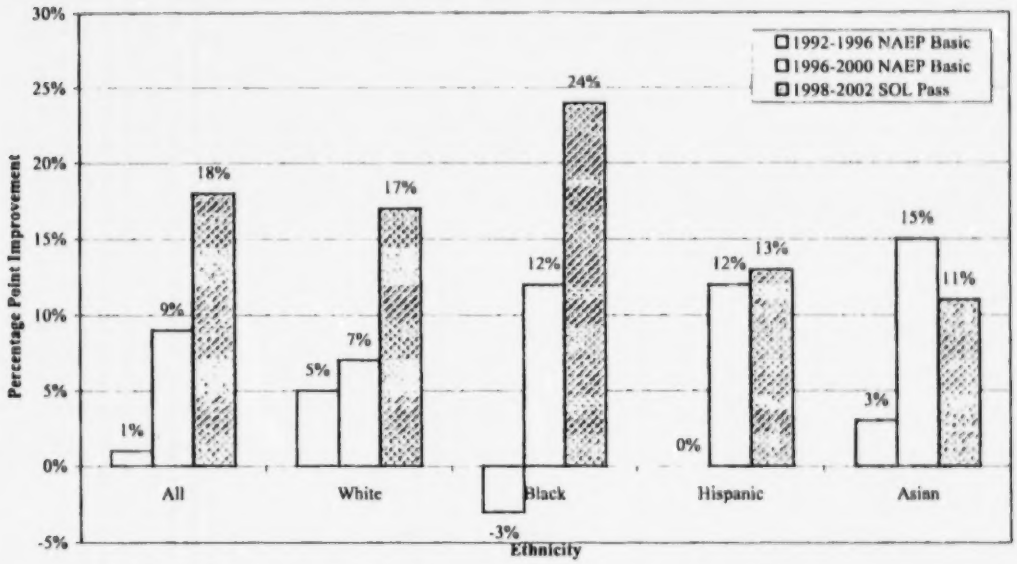
Question #2: What is a reasonable expectation for annual achievement gains, as measured by the NAEP basic level of performance?

Although NCLB does not require states to improve NAEP scores, both the House and Senate versions of NCLB set AYP targets based on the expectation that schools should increase "the percentage of students scoring at the proficient level or higher by at least one point per year" (Linn, Baker, & Betebenner, 2002). However, very few states, including Virginia and Georgia, were able to increase the percentage of students scoring at or above proficient on NAEP by one-percentage point a year. In many ways, this is not surprising since the NAEP proficient standard was set unrealistically high (Linn et al., 1996). Thus, the NAEP basic level may be a more attainable and realistic goal that states could attain with sufficient effort, resources, and time. Moreover, since there is substantial correspondence in the percentage of students meeting the state-defined proficiency standard and the NAEP basic level, it may be reasonable to expect states to show similar improvements on both assessments.

The purpose of the next set of descriptive analyses is to identify a reasonable expectation for improvement at the NAEP basic level and the state-defined proficiency standard on the Virginia SOL. To do this, we look at four-year trends at the NAEP basic level before and after the enactment of Virginia's SOL testing system in 1998. Since the SOL proficient standard corresponds to the NAEP basic level, we might expect stronger gains on the NAEP after 1996

since schools began to prepare students to do well on the SOL tests. And as shown earlier, basic performance on the NAEP appears strongly linked to proficient performance on the SOLs. Figure 5 displays four-year trends at NAEP basic and the Virginia SOLs in grade 8 math by

Figure 5: Four-Year NAEP Trends (1992 to 1996, 1996 to 20002) and Four-Year SOL Trends (1998-2002) by Racial and Ethnic Subgroups in Virginia



race and ethnicity. The solid bars show NAEP trends from 1992 to 1996, the dotted bars show NAEP trends from 1996 to 2000, and the striped bars show the four-year gains at the SOL-proficient level from 1998 to 2002. Before SOL testing, four-year (1992 to 1996) gains at the NAEP basic level were small for all students and subgroups. On average, there was only a 1-percentage point improvement for all students, and Black and Latino students showed either no annual gain or declines in performance. After SOL testing, the percentage of Black, Latino, and Asian student at or above NAEP basic increased by over 10-percentage points from 1996 to 2000. The magnitude of the four-year improvement at NAEP basic and the SOL proficient level was similar for Latino and Asian students. For Black and White students, however, SOL gains were about twice as large as NAEP gains.

BEST COPY AVAILABLE

Two important implications flow from these results. First, although is difficult to quantify the precise degree of correspondence to expect on NAEP with the SOL, it is certainly reasonable to expect that gains on the state test will be similar to gains at the NAEP basic level, as was the case for Latino and Asian students. Second, larger improvements at the NAEP basic level occurred after SOL testing. While there are a number of factors that may explain this finding, the overlap in content on NAEP and the SOLs may explain some of the gains on both assessments after 1996. In other words, the close correspondence in NAEP and SOL trends after the enactment of SOL testing suggests that we should expect some degree of correspondence in annual gains for both tests. But how large should the NAEP basic gains be for all students as well as racial and ethnic subgroups?

Table 4 (see next page) shows the increase at the NAEP basic level from 1990 to 2000 in selected Southern states. Virginia made at least a 1-point increase whereas Georgia fell short of that target. Among all states, North Carolina averaged a 3-point increase per year followed by three states—Texas, Kentucky, and West Virginia—which averaged a 2-point annual gain. Black and Latino students in most Southern states also met the 1-point criterion. In particular, ten-year gains for Blacks exceeded 20-percentage points in North Carolina and Texas, and ten-year gains were more than 20-percentage points for Latinos in Maryland, Texas, Virginia, and West Virginia.

Although the NAEP trends from the 1990s in selected Southern states suggest that a 2- to 3-point annual improvement on the NAEP basic level may be a realistic and ambitious target for states, many assumptions would have to be met in order for states to meet these achievement goals. First, the state NAEP gains in the 1990s occurred during a decade of economic prosperity, enabling many state governments to provide more money and resources for education.

Currently, however, most states are facing severe budget crises, and public schools will receive far fewer resources in trying to attain the ambitious 100% proficient target set by NCLB. Second, even though minority students in selected states such as Texas and North Carolina made impressive achievement gains (Grissmer & Flanagan, 1998), Black and Latino students made smaller gains in grade 8 math than Whites students in most states throughout the South (see Table 4). Thus, NAEP achievement trends from the 1990s suggest that it is unrealistic to expect all subgroups to make similar gains during the next 12-years. Third, Table 4 includes only achievement trends in math, which are generally larger than those in reading. For example, during a six-year year period in the 1990s (1990 to 1996), the percentage of students scoring at the NAEP basic level increased by 10-percentage points in grade 8 math but only 5-percentage points in grade 8 reading (Linn et al., 2002). Fourth, given the large racial and ethnic achievement gaps on most state tests, minority

Table 4: 10-Year Gains (1990-2000) at NAEP Basic in Grade 8 Math by Ethnicity.

State	All Students	White Students	Black Students	Latino Students
AL	12%	15%	6%	14%
AR	8%	10%	6%	9%
GA	8%	11%	5%	14%
KY	20%	20%	15%	n/a
LA	16%	26%	9%	12%
MD	15%	17%	13%	31%
NC	32%	33%	24%	47%
OK	12%	13%	13%	11%
TX	23%	19%	22%	30%
VA	15%	18%	12%	25%
WV	20%	20%	19%	27%

students will have to make much more rapid gains than Whites. For example, as shown in Figure 3 and 4 earlier, over 70% of White students met Virginia and Georgia's proficiency standard on the grade 8 math test in 2000. However, the proficiency rate was 38% for Blacks and 56% for Latinos in Virginia, and 30% for Blacks and 34% for Latinos in Georgia. Since

each subgroup is required to meet the proficiency standard within 12 years, Black and Latino students will have to make much larger annual gains than Whites. The magnitude of these mandated gains on a single state test will create strong incentives for schools to focus instruction narrowly on tested content. As a result, score gains on high-stakes state assessments may be inflated and may not lead to corresponding gains on low-stakes assessments that measure similar academic content (Koretz, Linn, Dunbar, & Shepard, 1990; Klein, Hamilton, McCaffrey, & Stecher, 2000). In future years, policymakers and researchers will need to know whether rapid gains on state tests represent genuine learning or inflated scores only on the state test. Just as NAEP is used to judge the rigor of state proficiency standards, policymakers may also rely on NAEP to verify large gains on state assessments and to assess whether score inflation poses a more serious among minority students than White students.

Question #3: What are the implications of the AYP racial subgroup requirements?

In addition to defining proficiency on the state test, NCLB requires state to establish AYP goals for all students and for each of the major racial and ethnic subgroups. Furthermore, to avoid federally mandated sanctions, schools must hit AYP targets for all students and for subgroups, since failure of any one subgroup to meet AYP targets means failure for the entire school. The AYP subgroup goals are intended to focus school efforts on improving the achievement of poor and minority students. However, AYP subgroup rules may also unfairly sanction racially diverse schools and exclude them from state-sponsored rewards programs for improved achievement. Since schools in the South and West are more racially diverse than schools in the North and Midwest, Kain and Staiger (2002a) point out that "the failure rate is likely to be two to four times higher in states in the South and West with large minority populations, because of the subgroup rules. Indeed, the single most important determinant of the

difference in failure rates between states is likely to be the racial composition of their schools" (p. 23).

There are two reasons why AYP subgroup rules put integrated schools at a statistical disadvantage. First, annual changes in test scores yield volatile estimates of mean school achievement due to differences in the sample of tested students and one-time factors like student misbehavior on the day of the test (Kane & Staiger, 2002b; Linn & Haug, 2002). Second, test score volatility is an even more serious problem for racial subgroups, since AYP targets will be based on smaller samples that yield unreliable averages. The statistically imprecise nature of test score changes has profound policy implications for racially diverse schools. As Kain and Staiger (2002) point out, requiring schools to meet several subgroup targets "is analogous to correctly calling 3 or 4 coin tosses in a row, instead of a single toss" (p. 258). In other words, integrated schools will have to hit more than one AYP target, and may be identified for school improvement simply because of random fluctuations in tests scores of any one subgroup.

Policy Recommendations

The descriptive analysis of achievement scores in Virginia and Georgia suggests that the state-defined proficiency level corresponds to the NAEP basic level. In other words, the NAEP basic level offers a minimum definition of proficiency on state performance standards. Linn, Baker, and Betebenner (2002) elaborate on the importance of using the NAEP basic level to judge the validity of score increases on the state assessments:

> If the percentages of students within each state who achieved at the basic level or higher on NAEP were used as a benchmark against which state standards of performance could be compared, it would assure that state standards were less disparate than they now are. At the very least, states having standards that had more students at the proficient level than at the basic level on NAEP might be required to provide a rationale to defend their levels (p. 15).

Because the NAEP basic level corresponds to the proficient level on the Virginia SOL and Georgia CRCT, it may be reasonable to expect similar and perhaps parallel improvements on both the state test and NAEP. We tested this hypothesis by exploring four-year trends at the NAEP basic level before and after the enactment of the Virginia SOL tests in 1998. We suspected that if the SOL proficient level was similar to NAEP basic level, four-year trends in NAEP basic would be larger after SOL testing since schools began to focus more heavily on content found in both assessments. In fact, annual gains at the NAEP basic level were larger after SOL testing started in 1998. Despite the correspondence between the NAEP basic level and Virginia and Georgia's proficiency definition on their state tests, there are many assumptions that would have to be meet for states to improve the performance of all subgroups. For example, it is unclear whether states will be able to accelerate the achievement of minority students at a faster rate than White students. Furthermore, most states are undergoing severe budget crises and may not be able to provide additional resources for helping schools meet NCLB's extremely ambitious achievement targets.

Discussion and Conclusion

During the first year of NCLB implementation, Virginia and Georgia encountered a unique set of opportunities and obstacles in complying with federal Title I requirements. The initial experience of both states suggests that collaboration across educational policymaking agencies will be critical if NCLB is to be implemented. For example, Virginia's accountability system operated in a stable political environment and many of its policies were already aligned with the principles and requirements underlying NCLB. In Georgia, however, political conflict between the governor and state superintendent of instruction produced education policies that were fragmented and incoherent. For example, testing requirements were expanded in grades 3

to 8 despite the need to upgrade the state curriculum on which assessments are based. This case study of Virginia and Georgia suggests that state capacity—that is, both collaboration among political and education leaders as well as the status of previous accountability policies—will shape policymakers' ability to implement NCLB's accountability requirements. It is important, then, to understand how federal Title I reforms operate in states with different education policies and governance arrangements. Based on these preliminary findings, we plan to conduct comparative analyses and in-depth case studies of NCLB implementation in four additional sites—Arizona, California, Illinois, and New York. These states differ in their governance arrangements in education and the extent to which previous policies conform to NCLB requirements.

In many ways, state implementation of NCLB's accountability requirements serves as a means to the primary objective of federal Title I policy, which is to bring all students to the proficiency level within 12 years. How high or low states define that proficiency level is likely to determine whether there is a realistic chance of all students reaching proficiency within 12 years. Since NCLB gives states considerable leeway in defining proficiency, it requires biennial NAEP testing to verify state proficiency definitions and annual achievement gains. The availability of NAEP data allows federal and state policymakers to address two questions related to the 12-year timeline for attaining 100% proficiency. First, how does the state definition of proficiency compare to NAEP performance levels? In Virginia and Georgia, the percentage of students scoring at the state-defined proficiency standard was virtually identical to the percentage at or above the NAEP basic level in grade 8 math. As a result, the NAEP basic level may be a minimum level for defining state proficiency standards.

Given the correspondence between the NAEP basic level and the state proficiency standard in Virginia and Georgia, policymakers also need to know: what is a reasonable expectation for annual achievement gains, as measured by the NAEP basic level of performance? Although several were to improve the percentage of students scoring at or above NAEP basic by at least 1-percentage point per year during the 1990s, it is unclear whether similar gains could be sustained over the next 12 years for all racial subgroups in both reading and math. As states endure increased financial hardship, reduced education spending may hamper local efforts to narrow the racial achievement gap. A third question raised by the federal law is whether all racial subgroups can meet adequate yearly progress objectives. Since there are large racial achievement gaps not only in Virginia and Georgia but also in most other U.S. states, minority students will have to make substantially larger gains than White students, creating strong pressures for schools to focus exclusively on tested content. Moreover, the accountability requirements in NCLB pose formidable challenges to racially diverse schools, which must meet multiple achievement targets. In many ways, racial equity underlies NCLB's stringent accountability provisions, underscoring the civil rights implications that are embedded in the federal law. Researchers and policymakers will have additional information in future years to determine whether the gains on state tests represent meaningful improvements in minority student achievement.

REFERENCES

American Federation of Teachers. (2001). *Making standards matter*. Washington, D.C.: American Federation of Teachers.

Associated Press. (2002, April 2). Governor reappoints DeMary to public schools post. *Associated Press State & Local Wire*.

Center on Education Policy. (2003). *From the capital to the classroom, state and federal efforts to implement the No Child Left Behind Act*. Washington, D.C.: Center on Education Policy.

Cohen, M. (2001). *Georgia assessment letter*. U. S. Department of Education, Office of Elementary and Secondary Education. Retrieved January 5, 2003, from the World Wide Web: http://www.ed.gov/offices/OESE/saa/ga.html

Consortium for Policy Research in Education. (2000a). *Assessment and accountability in the fifty states: 1999-2000 (Virginia)*. Philadelphia, PA: CPRE.

Consortium for Policy Research in Education. (2000b). *Assessment and accountability in the fifty states: 1999-2000 (Georgia)*. Philadelphia, PA: CPRE.

Education Week. (2002). *The changing definition of proficient.* Retrieved October 22, 2002, from the World Wide Web: http:www.edweek.org/ew/ew_printstory.cfm?slug=06tests-s2.h22

Georgia Board of Education. (2002). *Board reacts to superintendent's letter*. Georgia Board of Education. Retrieved February 28, 2002, from the World Wide Web: www.doe.k12.ga.us/board/newsletter/boe_news_jun_02.html

Georgia Department of Education. (2003). *Superintendent joins Governor in announcing S.T.A.R.S. education plan for Georgia*. Georgia Department of Education. Retrieved February 28, 2003, from the World Wide Web: www.doe.k12.ga.us/communications/releases/03/021803.html

Goertz, M. E., Duffy, M. D., & Le Floch, K. C. (2001). *Assessment and accountability systems in the 50 states: 1999-2000*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania, Graduate School of Education.

Grissmer, D., & Flanagan, A. (1998). *Exploring rapid achievement gains in North Carolina and Texas*. Washington, DC: National Education Goals Panel.

Hoff, D. J. (2002, October 9). *States revise the meaning of "proficient."* Education Week. Retrieved October 15, 2002, from the World Wide Web: http://www.edweek.org/ew/ew_printstory.cfm?slug=06tests.h22

Jacobson, L. (2003, March 4). Georgia's new leaders seek major governance changes. *Education Week*.

Kane, T. J., & Staiger, D. O. (2002a, June 10-11). *Racial subgroup rules in school accountability systems*. Paper presented at the Taking account of accountability: Assessing politics and policy, Kennedy School of Government, Harvard University.

Kane, T. J., & Staiger, D. O. (2002b). Volatility in school test scores: Implications for test-based accountability systems. In D. Ravitch (Ed.), *Brookings papers on educational policy, 2002* (pp. 235-283). Washington, DC: Brookings Institution Press.

Laccetti, S. (2000, October 13). Tactics have consequences: School board's petty games serve to humiliate Schrenko. *The Atlanta Journal and Constitution,* pp. 18A.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*, 4-16.

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher, 31*, 3-16.

Linn, R. L., & Haug, C. (2002). Stability of school-building accountability scores and gains. *Education Evaluation and Policy Analysis, 24*, 29-36.

Linn, R. L., Koretz, D. M., & Baker, E. L. (1996). *Assessing the validity of the National Assessment of Educational Progress: NAEP technical review panel white paper*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

O'Dell, L. (2002, August 24). SOLs have been revised enough. *The Associated Press State & Local Wire.*

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed. ed.). Newbury Park, CA: Sage.

Phi Delta Kappa International. (2002). *An external audit of the Georgia Quality Core Curriculum*. Bloomington, IN: International Curriculum Management Audit Center, Phi Delta Kappa International.

Portner, J. (1999, January 20). Massive failure rates on new tests daze VA. *Education Week.*

Ravitch, D. (2002). Introduction. In D. Ravitch (Ed.), *Brookings papers on education policy 2002* (pp. 1-12). Washington, D.C.: Brookings Institution.

Robelen, E. W. (2002, April 17). *States, Ed. Dept. reach accords on the 1994 ESEA*. Education Week. Retrieved September 17, 2002, from the World Wide Web: http://www.edweek.org/ew/ew_printstory.cfm?slug=31esea.h21

Salzer, J. (2003). Perdue names 4 to seats on school board; redistricting rule lets him dump some holdouts. *The Atlanta Journal and Constitution,* pp. 1C.

The Virginian-Pilot. (2002, April 5). Warner taps steady hand at time of school turmoil. *The Virginian-Pilot.*

Virginia Board of Education. (2000). *Regulations establishing standards for accrediting public schools in Virginia*. Richmond, VA: Virginia Board of Education.

Virginia Board of Education. (2002). *Committee to implement the "No Child Left Behind Act"*. Commonwealth of Virginia, Board of Education. Retrieved March 2, 2003, from the World Wide Web: http://www.pen.k12.va.us/VDOE/meetings/2002/9-04NCLB.pdf

Virginia Department of Education. (2002). *Virginia's consolidated state application*. Richmond, VA: Virginia Department of Education.

Wirt, F. M. (1977). School policy culture and state decentralization. In J. D. Scribner (Ed.), *The politics of education, The seventy-sixth yearbook of the National Society for the Study of Education, part 2* (pp. 164-187). Chicago: University of Chicago Press.

TM035124

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

Reproduction Release

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: The initial response to the accountability requirements in the No Child Left Behind Act: A case study of Virginia and Georgia	
Author(s): Jimmy Kim	
Corporate Source: The Civil Rights Project, Harvard University	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents

Level 1	Level 2A	Level 2B
X		
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only
Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.		

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature:	Printed Name/Position/Title: Jimmy Kim, Research Associate	
Organization/Address: The Civil Rights Project 125 Mt. Auburn Street, 3rd Floor Cambridge, MA 02138	Telephone: 617-495-3617	Fax: 617-495-5210
	E-mail Address: jimmykim@law.harvard.edu	Date: 6-6-03

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706
Telephone: 301-552-4200
Toll Free: 800-799-3742
e-mail: ericfac@inet.ed.gov
WWW: http://ericfacility.org

EFF-088 (Rev. 2/2001)